DOCUMENT RESUME

ED 119 427 EC 081 557

AUTHOR Rudner, Lawrence M.
TITLE The Development and Evaluation of a Quasi Non Verbal Cognitive Skills Test.
SPONS AGENCY Department of Health, Education, and Welfare, Washington, D.C.
PUB DATE Mar 76
NOTE 31p.; Some pages in appendix may not reproduce clearly due to quality of original document

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage
DESCRIPTORS *Aurally Handicapped; Cognitive Development; *Cognitive Measurement; Comprehension; Exceptional Child Research; Maturation; *Nonverbal Tests; Secondary Education; Test Construction; *Test Reliability; *Test Validity

ABSTRACT

Summarized are the development and evaluation of a quasi non-verbal Cognitive Skills Test (CST) to measure comprehension, application, analysis, and synthesis in hearing impaired adolescents. It is explained that the measure was evaluated with 113 students at a school for the deaf over a 2-year period. Among findings reported are that the test and each subtest reflect a high degree of internal consistency, but that all the subtests may be measuring the same trait (perhaps general intelligence) rather than the target cognitive skills. It is recommended that the CST be used for research purposes only. The entire test is appended. (LS)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED119427

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Development and Evaluation of a Quasi Non Verbal Cognitive Skills Test

Lawrence M. Rudner

Model Secondary School for the Deaf
Gallaudet College
Washington, D.C.

March, 1976

Produced with funds from the US Department of Health, Education and Welfare, PL-89-694. However the opinions expressed herein do not necessarily reflect those at the US Department of H.E.W.

EC 081 557

This paper summarizes the development and evaluation of a quasi non-verbal Cognitive Skills Test (CST) designed to measure competency levels in the cognitive developmental areas of comprehension, application, analysis and synthesis as defined by Bloom, et. al. (1956). Unlike existing measures this cognitive skills test 1) is not directed toward any one particular subject area and 2) is designed specifically for use with hearing impaired adolescents.

The CST, and a verbal counterpart, were designed to meet an in-house need during the evaluation of an experimental two year program (see Chewning, 1973). As such the CST is intended to serve solely as a potential indicator of the above skills, not as a definative measure. The CST has not been fully validated nor has it been normed for any population. Consequently the CST should be used for research purposes only.

Description of the Measure

The CST is composed of four, 10 item subtests designed to measure the following cognitive skills (descriptions are rewordings of Bloom taxonomy by Gronlund, 1970):

COMPREHENSION -- "The ability to grasp the meaning of material. This may be shown by translating material from one form to another (words to numbers), by interpreting material (explaining or summarizing), and by estimating future trends (predicing consequences or efforts). These learning outcomes go one step beyong the simple remembering of material."

APPLICATION -- "The ability to use learned material in new and concrete situations. This may include the application of such things as rules, methods, concepts, principles, laws, and theories."

ANALYSIS -- "The ability to break down material into its component parts so that its organizational structure may be understood. This may include the identification of the parts, analysis of the relationships between parts, and recognition of the organizational principles involved . Learning outcomes here require an understanding of both the content and the structural form of the material."

SYNTHESIS -- "The ability to put parts together to form a new whole. This may involve the production of a unique communication (theme or speech), a plan of operations (research proposal), or a set of abstract relations (scheme for classifying information). Learning outcomes in this area stress creative behaviors, with major emphasis on the formulation of new patterns or structures."

The CST is quasi non-verbal in that the alternate responses are largely pictorial. While item directions are usually written, only minimal reading skills are involved. To further minimize the neccesity of reading skills, the test administrator is encouraged to explain the item directions to all examinees when they are not fully comprehended by any examinee.

Development of the CST

An initial item pool of 67 items was developed. Each item was designed such that the requisite skill matched only one of the classification areas outlined by Bloom. Content material was selected so as to be appropriate for an adolescent hearing impaired population.

This initial item pool was then presented to a jury of five deaf and hearing research and evaluation specialists having experience with Bloom's taxonomy, deaf education and test devlopment. Jury

members were asked to judge which cognitive skill was reflected by each item and whether item content was clear and appropriate for the intended population.

Fifty-seven items achieving a high degree of consensus were retained for field testing. Several of the drawings were altered to enhance clarity. Following a reccomendation by Guilford (1954) additional distractors were added to several of the multiple choice items.

The pool of 57 items was then field tested with 20 high school students at the Lexington School for the Deaf in New York City, through the courtesy of their research department. These students varied in chronological age, scholastic success and number of years in high school. After the testing session some of the students discussed the CST with the designer. Ideas were collected on improving and altering the measure. Questions and drawings which caused undue difficulty and/or lacked clarity were identified for future modification.

This subjective information as well as each item's discrimination and difficulty indices, were considered in developing the present version of the CST. Typically the ten items within each subtest having the highest discrimination value were retained. Variations insured a wider range of item difficulties and the omission of confusing or misleading items. Poor distractors were either dropped or modified. The resultant 40 item measure (See Appendix) was then used for research purposes at the Model Secondary School for the Deaf (MSSD).

Evaluation of the Measure

The following analysis is based on the item responses made by 113 students at the MSSD over a two year period.

As shown in Table 1, the test and each subtest reflect a high degree of internal consistancy. The 100 item Kuder Richardson Formula 20 reliabilities are above .90. The Standard Errors of Measurement for the subtest and the battery are within reasonable limits (see Rudner, 1973). A perusal of table 2 shows that the items within each subtest reflect a fair variety of item difficulties (MEAN) and fairly high discrimination indices [R (SCALE)]. Table 3 shows that the erroneous responses were fairly well distributed over the distractors. However, several poor distractors are presented.

To obtain some insight into the validity of the CST, each subtest was subjected to a one factor principle component analysis. Following the recommendation of Maxwell (1959) and Tobias and Carlson (1969), Barlett's test of sphericity (1950) was applied to each subtest inter-item tetrachoric correlation matrix. This test has the associated hypothesis that the sample correlation matrix comes from a multivariate normal population in which the items are uncorrelated. (See Cooley and Lohnes, 1971). That is, this test was used to ascertain the appropriateness of using a factor analytic approach. As shown in Table 4, this hypothesis was rejected at $p < .05$ for the matrices from each subtest. While the hypothesis could also be rejected for the tetrachoric correlation matrix of all the items within the battery, it was felt that the N = 113 subjects did not warrant a factor analysis of the entire battery.

Table 1

Baseline Statistics for the CST and Constituant Subtests[1].

	Subtests				
	Comprehension	Application	Analysis	Synthesis	Total Battery
# of items	10	10	10	10	40
Optimum Mean[2]	6.08	6.58	6.48	5.86	25.00
Mean	7.58	7.16	6.42	4.92	26.10
Standard deviation	1.64	1.97	1.88	2.08	6.24
Reliability (KR-20)	.5028	.5891	.4966	.5774	.8313
100 item KR-20	.9100	.9348	.9080	.9318	.9250
Standard Error of Measurement	1.15	1.26	1.33	1.35	2.56

1. based on N= 113 examinees

2. Mean for highest possible KR-20, considering the varying number as response alternatives

TABLE 2

ITEM STATISTICS

ITEM	SCALE	MEAN	SIGMA	R(TOTAL)	R(SCALE)
1	1	0.90	0.29o	0.1922	0.2075
2	2	0.81	0.389	0.22o9	0.3398
3	3	0.o5	0.475	0.3166	0.3413
4	2	0.53	0.499	0.4792	0.4o56
5	3	0.58	0.494	0.2900	0.4975
6	2	0.65	0.475	0.3285	0.4477
7	4	0.51	0.500	0.49o0	0.5347
8	3	0.88	0.319	0.1572	0.34o1
9	2	0.80	0.403	0.3855	0.4987
10	3	0.58	0.494	0.43o3	0.4785
11	4	0.38	0.486	0.2116	0.34o6
12	4	0.91	0.284	0.2o49	0.3473
13	3	0.32	0.46o	0.1405	0.2590
14	2	0.43	0.49o	0.2112	0.3501
15	3	0.96	0.185	0.1414	0.1194
16	2	0.80	0.403	0.4101	0.4320
17	2	0.o3	0.483	0.4210	0.5099
18	1	0.96	0.20o	0.1898	0.2335
19	1	0.40	0.490	0.4438	0.5137
20	3	0.73	0.446	0.3378	0.4752
21	2	0.82	0.382	0.4909	0.5437
22	3	0.73	0.44o	0.3569	0.4436
23	4	0.42	0.494	0.53o2	0.5260
24	4	0.31	0.4o2	0.54o8	0.6547
25	4	0.o4	0.481	0.5585	0.5998
2o	1	0.81	0.389	0.2998	0.4880
27	1	0.50	0.500	0.2980	0.4338
28	4	0.27	0.442	0.3976	0.5095
29	4	0.73	0.442	0.19o5	0.3057
30	3	0.51	0.500	0.5046	0.5293
31	1	0.93	0.25o	0.2811	0.4128
32	1	0.74	0.437	0.39o4	0.4306
33	1	0.96	0.206	0.2588	0.2597
34	1	0.78	0.415	0.4o34	0.4487
35	2	0.83	0.374	0.5840	0.o246
3o	2	0.86	0.349	0.4665	0.4710
37	3	0.49	0.500	0.55o5	0.6258
38	1	0.o1	0.488	0.5691	0.7032
39	4	0.28	0.451	0.1246	0.24o9
40	4	0.47	0.499	0.3818	0.4731

Table 3

Percentage of students Selecting each alternative response in Multiple Choice Items[1]

Item	1	2	3	4	5
1	4.4	1.8	*	0.9	0.0
2	11.5	2.7	2.7	*	-
3	*	5.3	15.0	14.1	-
4	4.4	4.4	29.2	*	0.9
5	26.5	5.3	*	6.2	-
6	2.6	4.4	1.8	*	23.0
7	*	3.5	3.5	0.8	39.8
8	0.0	*	0.9	1.7	7.1
9	3.5	2.7	*	13.3	-
10	12.3	19.5	4.4	*	-
11	22.1	31.9	*	6.2	-
12	26.7	1.8	1.8	*	.9
13	23.9	8.8	*	4.4	21.2
14	23.0	15.0	4.4	*	-
15	0.0	2.7	0.0	*	0.9
16	9.7	1.8	8.0	*	-
17	26.5	1.8	8.0	*	-
18	0.9	0.0	0.0	0.9	*
19	*	6.2	15.9	31.9	-
20	15.0	*	5.3	1.8	0.9
31	2.7	8	1.8	1.8	-
32	15.0	4.4	4.4	*	-
33	0.9	0.0	*	3.5	-
34	*	11.5	7.1	2.7	-
39	12.4	*	15.9	28.3	-

1. Based on N=113 subjects. *indicates correct response - indicates alternative not presented

Unlike other methods of factor analysis, principle component analysis provides factors which are directly obtained from the item responses (rather than implied by the data). When the number of factors is set equal to one, the factor loadings can be interpreted as the correlation between each item and the trait (i.e., dimension or factor) defined by the subtest. In a truly unidimensional subtest, each item would reflect a substantial item-trait correlation. As shown in Table 5, this was partially confirmed. Within each subtest, most items appear to be measuring the same trait. That is, within each subtest, the factor loadings of each item was fair to substantial ($r > .25$). It should be noted however, that there are a substantial number of items failing to meet this criteria.

Conclusions and Reccommendation

While the CST and its constituant subtest appear to be internally consistant and each subtest appears to be somewhat unidimensional, one cannot say with assurity that the CST actually measures the cognitive skills of comprehension, application, analysis and synthesis. It is entirely possible that all the subtests measure the same trait -- perhaps some form of general intelligence.

Until correlative data can be obtained, the validity of the measure must be questioned. A proper validation study would involve obtaining definative data which pinpoints each student's competency level within each of the assessed cognitive areas. In as much as most classroom learning is geared to the lower levels and since such classification would be extremely difficult, this type of formal validation may not be economically feasible.

Table 4

The Bartlett Test of
Sphericity for CST and subtests

	Subjects	Items	df	x^2	p
Comprehension	113	10	45	105.74	$<.05$
Application	113	10	45	150.46	$<.05$
Analysis	113	10	45	70.37	$<.05$
Synthesis	113	10	45	115.66	$<.05$
Battery	113	40	780	1325.90	$<.05$

Table 5

Factor Loadings
Using separate Principle Component
Factor Analyses

Comprehension Subtest		Application Subtest		Analysis Subtest		Synthesis Subtest	
Item	Loadings	Item	Loadings	Item	Loadings	Item	Loadings
1	.125	2	.202	3	.242	7	.530
18	.235	4	.321	5	.536	11	.117
19	.533	6	.404	8	.376	12	.384
26	.540	9	.479	10	.447	23	.527
27	.370	14	.174	13	.042	24	.761
31	.454	16	.406	15	.090	25	.696
32	.372	17	.488	20	.536	28	.596
33	.281	21	.610	22	.472	29	.146
34	.413	35	.781	30	.569	39	.011
38	.754	36	.663	37	.655	40	.482

It is therefore recommended that the CST be used for research purposes only. The researcher may wish to combine subtest scores and use the measure as a rough indicator of cognitive compentency. The ten items within each subtest cannot provide five discriminations between students. One potential use is as a pre-post measure over a fairly long time span. The researcher may also wish to compare raw scores with the means and standard deviations provided in Table 1. However one must recall that these statistics are based on students in one school for the deaf and hence should not be treated as representative of population values.

References

Bartlett, M.S. Tests of Significance in Factor Factor Analysis. British Journal of Psychology, Statistical Section, 1950, 3, 77-83.

Bloom, B. et. al Taxonomy of Educational Objectives The Classification of Educational Goals, Handbook I : Cognitive Domain, New York, McKay, 1956.

Chewning, B. A proposal for the establishment of a mini school. Unpublished manuscript, 1973.

Cooley, W.W. and Lohnes, P.R. Multivariate Data Analysis, New York; John Wiley & sons, 1971.

Gronlund, N. Stating Behavioral Objectives for Classroom Instruction, New York; MacMillian, 1970.

Guilford, J.P. Psychometric Methods, New York; McGraw-Hill, 1954.

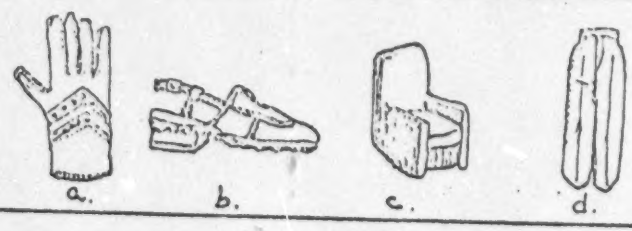
Maxwell, A.F. Statistical Methods in Factor Analysis, Psychological Bulletin, 1959, 56, 228-235.

Rudner, L.M. Simplified Item Analysis, Danbury, Conn.; Western Connecticut State College Press, 1973.

Tobias, S. and Carlson, J.E. Brief report: Bartlett's Test of sphericity and chance findings in factor analysis, Multivariate Behavioral Research, 1969, 4, 375-377.

APPENDIX

Sample Questions

1. Which does not belong?	 a. b. c. d.
2. 1,2,3,4,__?__,__?__	____, ____

In sample question 1, you should have circled choice c.

1. Which does not belong?	

You can wear a mitten, a shoe or pants, but you cannot wear a chair. The chair does not belong with the others and should have been circled.

In sample question 2, you should have written the numbers 5 and 6 in the blanks.

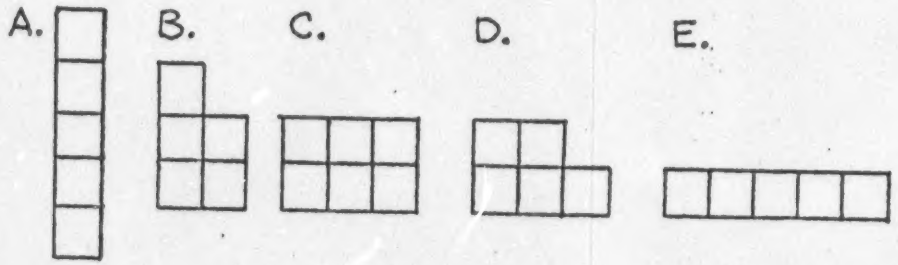
2. 1,2,3,4,__?__,__?__	5, 6

1, 2, 3, 4 is just counting. The next numbers are 5 and 6.

Name _____

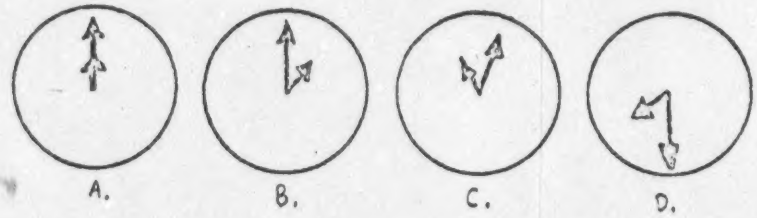
1.

Which has the most?



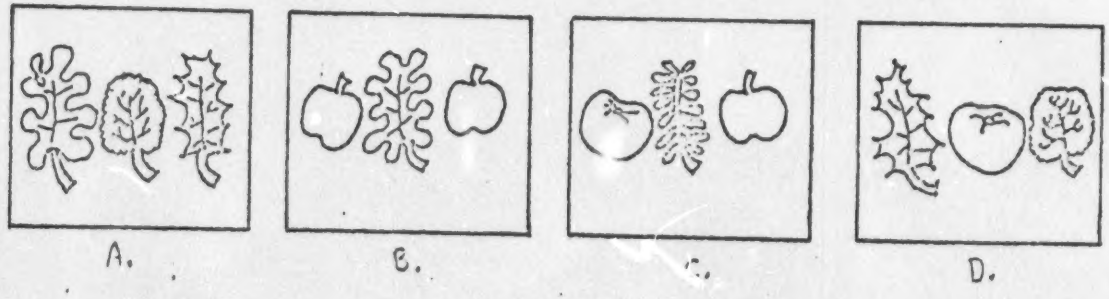
2.

What time is it?



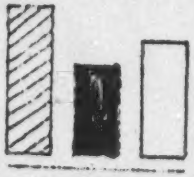
3.

Which one does not belong with the others?

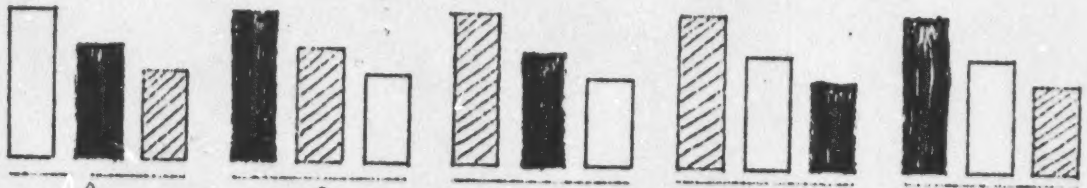


Name _____ page 2

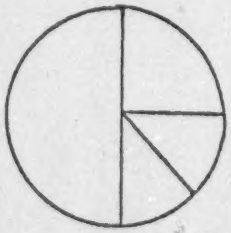
4.



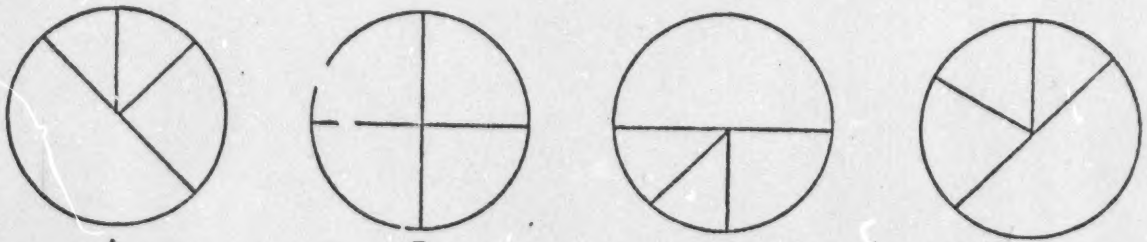
Put these in order.



5.

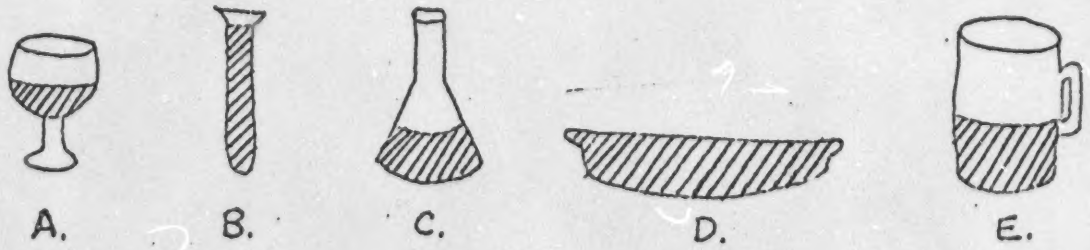


Which is the same?



6.

Which has the most water?

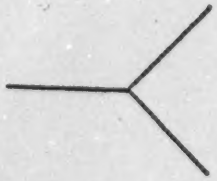


Name _____ page 3

7.

A. B. C. D. no answer E.

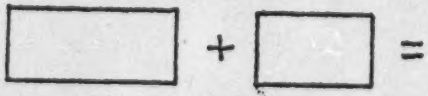
8.



Which is the same?

A. B. C. D. no answer E.

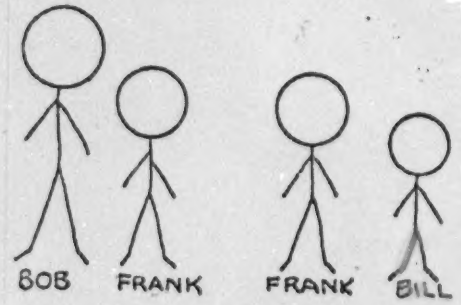
9.

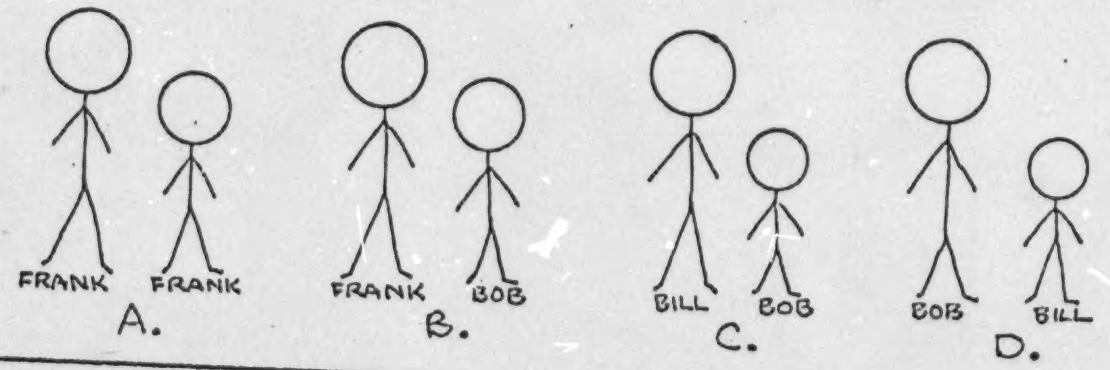


A. B. C. D.

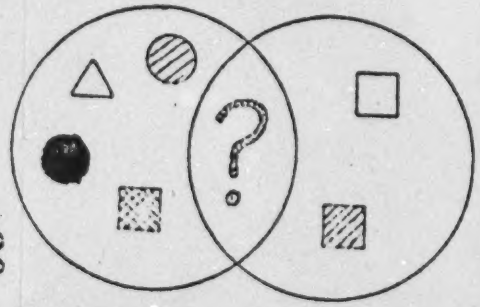
Name _____

10.



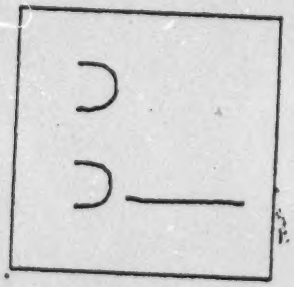


11. Which goes here?



20

12. What letter is this?



Name _____

13.

X + Y =

X =

Y = ____

A.

B.

C.

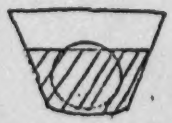
D.

no answer
E.

14.

21

A.


B.

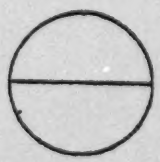
C.

D.

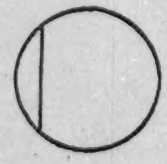
15

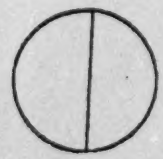
Which does not belong?

A.


B.

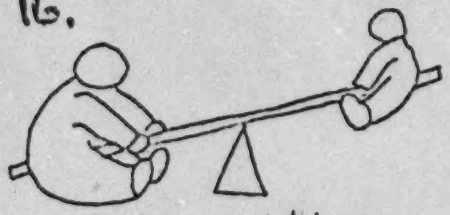
C.


D.

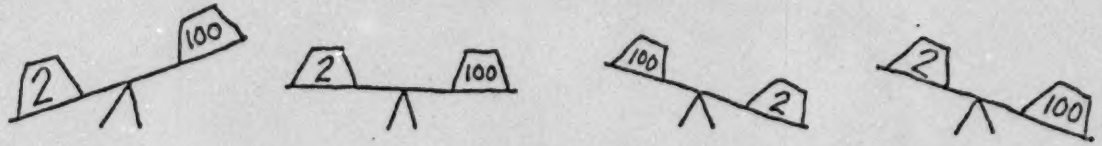

E.

Name _____ page 6

16.



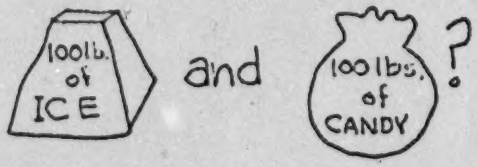
What would happen with a [100] lb. weight and a [2] lb. weight?



A. B. C. D.

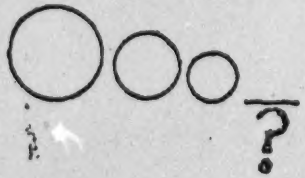
17.

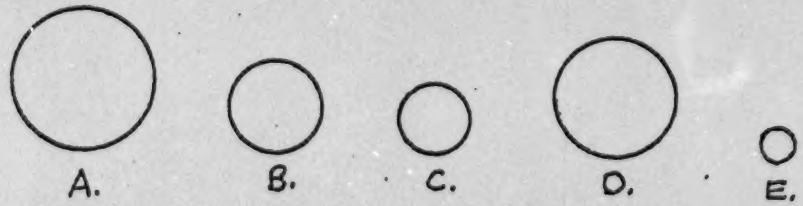
What would happen with



A. B. C. D.

18.



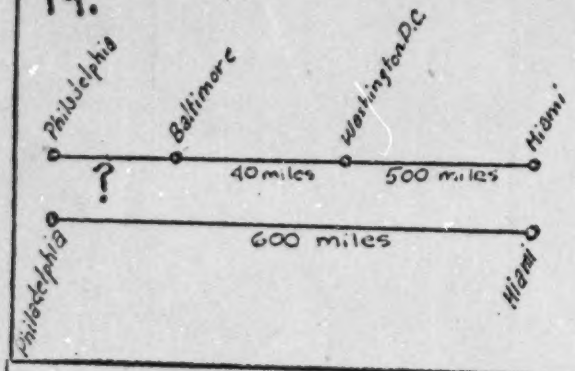


A. B. C. D. E.

22

Name _____

19.



A. 600 miles – (40 miles + 500 miles)

B. 600 miles – 500 miles

C. 500 miles + 40 miles

D. 600 miles + 500 miles + 40 miles

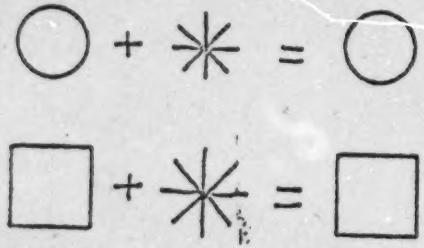
20.

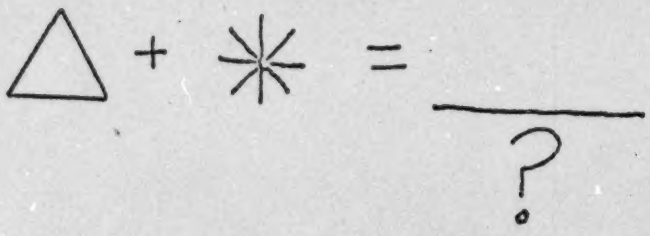
A, C, E, G, ?, K, M

H I J F L

A. B. C. D. E.

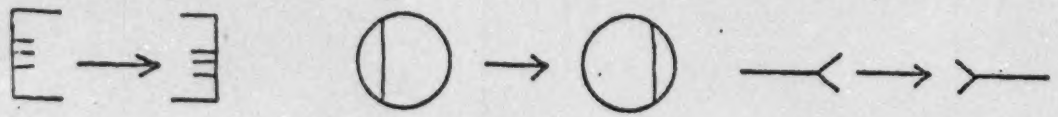
21.

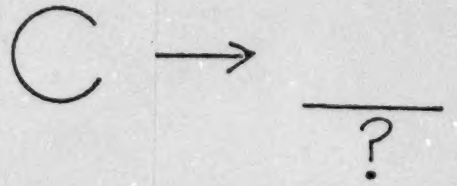




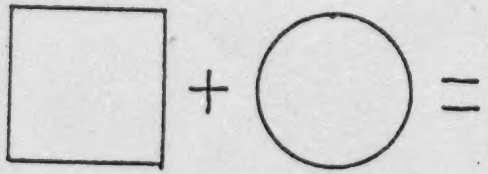
Name _____ page 8

22.



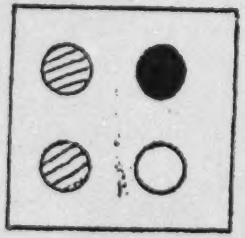


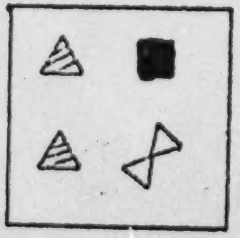
23.



Draw your answer

24. Here are two groups.



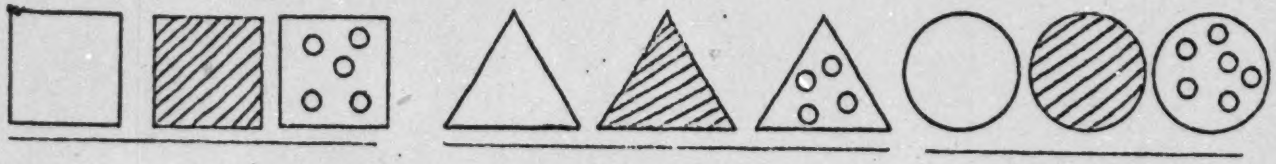


Make one more yourself.

24

Name _____ page 9

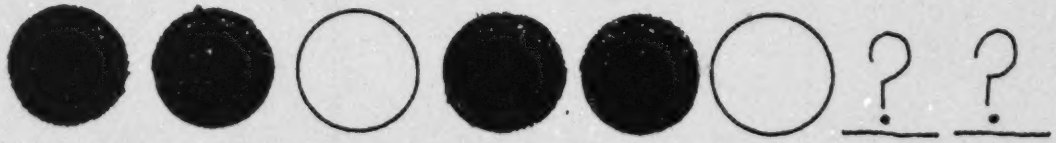
25.



Make another group.

_____

26.



_____ _____

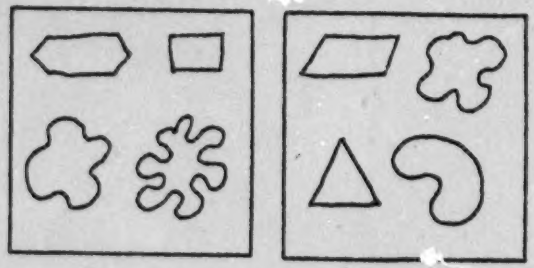
27.

1, 2, 4, 5, 7, 8, ?

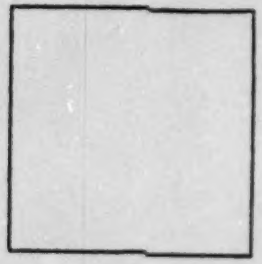
_____

Name _____ page 10

28.



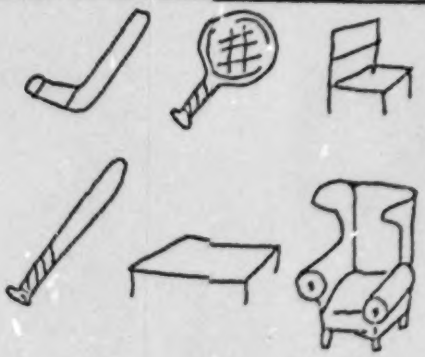
Make a new group using 2 things from each old group.



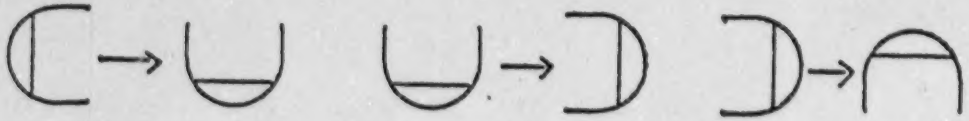
29.

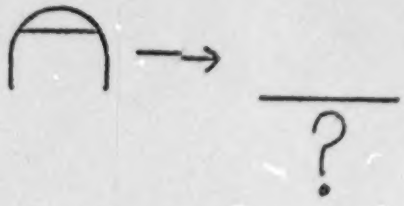
Circle the things that go together.

26



30.





Name _____ page 11

31. How many people lived in Newcity in 1968?

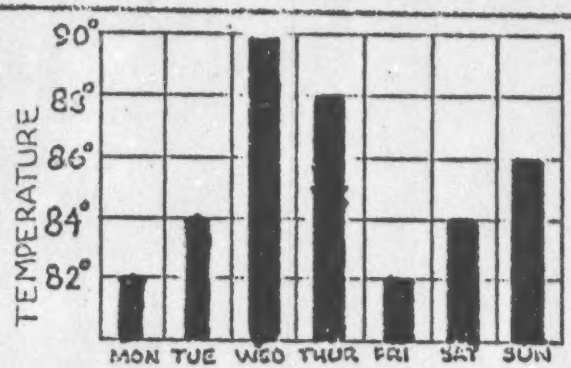
27

A. 50 B. 75 C. 100 D. 150

32. There were more people in Newcity in 1969 than in _____?

A. 1970 B. 1971 C. 1972 D. 1968

Name _____ page 12



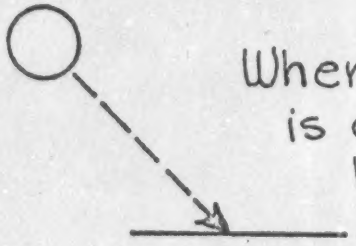
33. What was the temperature on Saturday?

A. 80° B. 82° C. 84° D. 86°

34. Saturday was hotter than _____.

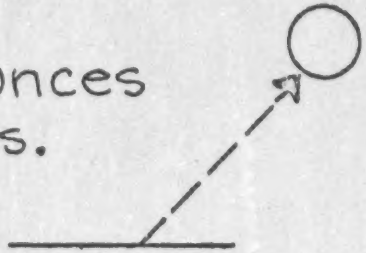
A. Monday B. Sunday C. Thursday D. Tuesday

Name _____ page 13



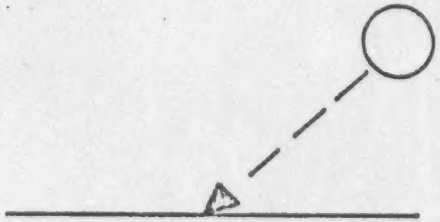
When a ball is dropped like this,

it bounces like this.



35.

When a ball is dropped like this,



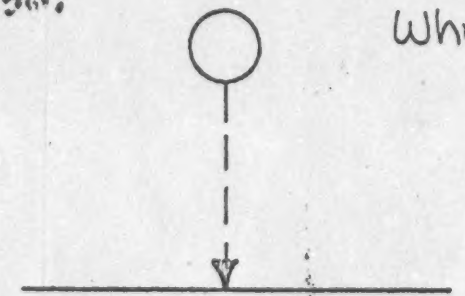
how does it bounce?

_____

?

36.

When a ball is dropped like this,



how does it bounce?

_____

?

Name _____

31.

1, 2, 5, 6, ?, ?, 13, 14

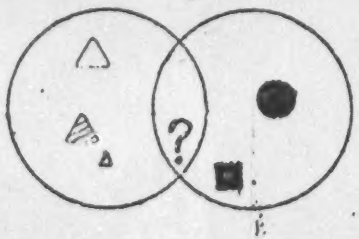
_____ _____

38. What comes next?

?

_____

37. Which goes here?



A. B. C. D.

Name _____ page 15.

40.

Put these in order from smallest to largest.